Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Gender Differences in the Adoption of Prompt Engineering in Generative AI.

Dr. Nubi Achebo [1]
[1]*Subject Matter Expert,* Nigerian University of Technology

ARTICLE INFO

Received: 08 Nov 2024

Revised: 18 Dec 2024

Accepted: 28 Dec 2024

ABSTRACT

Introduction: Generative artificial intelligence (AI) has rapidly reshaped digital work and learning, and prompt engineering the deliberate design and iterative optimisation of inputs to generative models has emerged as a key, practice-oriented competency. Concurrently, substantial literature on digital skills and AI adoption documents persistent gendered disparities in access, participation and confidence. Yet, systematic synthesis focused specifically on gender differences in the adoption of prompt engineering remains limited. This study synthesises existing scholarship, industry reports and platform data to characterise the state of knowledge and to identify levers for more equitable AI literacy.

Objectives: To (1) review and synthesise literature on the concept and evolution of prompt engineering; (2) examine studies and reports on gender differences in digital technology and AI adoption; (3) analyse secondary datasets and industry surveys to infer patterns in awareness and adoption of prompt engineering across genders; (4) evaluate the applicability of TAM and UTAUT theoretical frameworks to prompt-engineering adoption; (5) identify critical factors and barriers that shape gender disparities; and (6) propose evidence-based strategies for promoting gender-inclusive AI literacy.

Methods: A systematic secondary-research approach was employed. Searches were conducted across academic databases and institutional repositories for literature dated 2015, using targeted keywords (e.g., “prompt engineering,” “generative AI,” “gender AND AI adoption,” “TAM,” “UTAUT”). Inclusion criteria prioritised peer-reviewed studies, reputable industry reports and policy documents. Extracted materials were coded into thematic categories and analysed through qualitative content analysis, thematic synthesis and comparative evaluation, with interpretive mapping onto TAM/UTAUT constructs.

Results: The synthesis shows that prompt engineering is best characterised as a socio-technical, transdisciplinary skill combining domain knowledge, communicative framing and iterative model-testing. Secondary evidence indicates gendered patterns: women are under-represented in many GenAI technical courses and AI engineering roles (platform analytics and workforce reports suggest female shares near one-third in many GenAI enrollments), report higher AI-related anxiety and lower self-efficacy, and have less exposure to developer communities that foster tacit learning. Triangulated inferences identify (a) an exposure differential favoring men; (b) motivational divergence—men respond strongly to performance expectancy while women are more sensitive to effort expectancy and social facilitation;

and (c) design sensitivity scaffolded curricula, no-code tools, cohort mentorship and domain-relevant modules disproportionately increase female participation. Major barriers include access/infrastructure gaps, educational pipeline deficits, time poverty, mentorship shortfalls and ethics/trust concerns. Recommended enablers include scaffolded entry points, contextualised micro-credentials, mentorship/cohort models, institutional supports (devices, time-flexible training, gender-disaggregated monitoring) and embedded ethics training. The review also identifies key gaps: objective, gender-disaggregated measures of prompt-engineering competence and longitudinal evaluations of training-to-outcome pathways are scarce.

Conclusions: Prompt engineering is a teachable and high-value competency whose equitable diffusion depends on addressing structural, pedagogical and perceptual constraints. Interventions that reduce effort expectancy, strengthen social facilitation and increase facilitating conditions are theoretically and empirically justified to narrow gender gaps. To move from inference to evidence, future research should develop validated prompt-quality metrics, collect representative gender-disaggregated outcome data (including non-binary categories), and evaluate interventions via longitudinal and mixed-method designs.

Keywords: Prompt Engineering, Generative AI

INTRODUCTION

Generative Artificial Intelligence (GenAI) refers to a class of machine-learning systems capable of producing entirely new data whether text, images, video, audio, or code based on patterns learned from massive datasets. Popular examples include OpenAI's GPT models, Google's Gemini, Anthropic's Claude, Stability AI's Stable Diffusion, and Midjourney. Unlike discriminative AI systems, which primarily classify or predict outcomes, generative models synthesize new content conditioned on prompts or contexts provided by users. Underpinning these systems are advanced neural architectures such as transformers, which rely on self-attention mechanisms to model long-range dependencies in sequences. Large language models like GPT-4 or GPT-5 employ billions of parameters and immense datasets to acquire probabilistic representations of language. Text-to-image systems such as Stable Diffusion or DALL-E combine text encoders with diffusion models or generative adversarial networks (GANs) to translate text descriptions into images. In every case, these architectures depend critically on conditioning signals namely, the prompt or input supplied by the user which determines the nature, style, and quality of the generated output. Prompt engineering has become essential due to model complexity, cost efficiency, and customization potential. Modern large language models and diffusion models contain billions of parameters whose behavior is emergent and non-deterministic, meaning small changes in prompt structure or parameters can produce disproportionately large changes in the generated output. By crafting well-designed prompts, users can drastically reduce the number of iterations or API calls required to obtain a usable result, thereby saving computational resources, latency, and cost. In addition, prompt engineering enables domain-specific customization without the expense of retraining models from scratch. By embedding domain-rich prompts, enterprises and individual users can elicit near-custom behavior from off-the-shelf generative models, effectively adapting them to specialized tasks, industries, or knowledge bases. Although gender is a sociocultural construct, it correlates with distinct patterns of technology use and skill acquisition, and from a technical standpoint these differences shape how individuals interact with and optimize AI systems. Access to hardware and connectivity can vary, and differential access to high-performance devices or

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

broadband networks can limit opportunities to experiment with GenAI tools. Exposure to technical education, particularly in STEM disciplines, affects familiarity with algorithmic thinking, coding, and digital tools all of which are directly relevant to prompt engineering. Confidence and self-efficacy also play a role: technically, lower confidence can lead to less experimentation with parameters, advanced prompting methods, or API integrations. Variations in learning styles influence the preferred types of resources such as tutorials, online communities, or formal documentation that each gender may rely on to internalize prompt engineering practices. Even risk tolerance in adjusting parameters matters, since some prompt engineering tasks such as increasing temperature to explore creative extremes or testing boundary prompts require a willingness to experiment with potentially unpredictable outputs.

LITERATURE REVIEW

Brown (2020) introduced GPT-3, demonstrating that massively scaled autoregressive language models (175B parameters) exhibit strong few-shot capabilities: they can perform new tasks by conditioning on a handful of examples in the prompt rather than requiring parameter fine-tuning. Brown et al. systematically evaluated task performance across multiple NLP benchmarks and highlighted how model behavior is highly sensitive to prompt design, example selection, and ordering. Technically, the paper established that the prompt functions as an implicit task specification and that prompt engineering can elicit complex behavior from a single, pre-trained model. This finding underpins later technical work showing that well-constructed prompts (few-shot exemplars, instruction clarity, etc.) can approximate task-adapted models, which has direct implications for adoption: users who can craft effective few-shot prompts can obtain domain-specific performance without model retraining. The GPT-3 work also made clear the role of scale in emergent capabilities, cementing prompt engineering as a practical way to steer large models an insight central to modern prompt engineering toolchains and evaluation frameworks.

Wei et al. (2022) introduced and empirically validated chain-of-thought (CoT) prompting, a technique that asks language models to produce intermediate reasoning steps before giving a final answer. They showed that providing a few exemplars with explicit reasoning chains dramatically improves performance on arithmetic, commonsense, and symbolic reasoning tasks for sufficiently large models. Technically, CoT changed prompt engineering from single-step instruction to structured multi-step prompting, revealing that how a prompt solicits internal reasoning alters model latent trajectories and final outputs. The paper also exposed practical constraints CoT benefits appear primarily with larger models and appropriate exemplars, and they increase token usage and cost. For practitioners, CoT demonstrates an advanced prompting family that boosts reasoning but requires more sophisticated prompt templates, careful example selection, and cost-aware design skills that separate novice users from advanced prompt engineers. This has implications for adoption: users must understand model scale, prompt structure, and cost trade-offs to effectively apply CoT techniques.

Lewis et al. (2020) proposed combining parametric LMs with a non-parametric retriever to ground generation in external documents, improving factuality and reducing hallucinations. The RAG framework conditions generation on retrieved passages and jointly trains retrieval and generation components, which materially affects prompt engineering: instead of relying solely on prompt context, users (or systems) can programmatically augment prompts with retrieved knowledge snippets, shifting part of the engineering effort to retrieval pipelines, indexing, and prompt templates that incorporate documents. Technically, RAG shows that effective prompt engineering can involve building retrieval-to-prompt workflows, token-budget management for concatenated context, and template design that blends retrieved evidence with user instructions. Adoption of such techniques requires additional technical competencies vector search, document chunking, and prompt templating beyond simple prompt phrasing. Thus, RAG both expands the toolkit of prompt engineering and raises the bar for users who wish to achieve high factual accuracy in downstream applications.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Schulhoff (2024) synthesized the rapid proliferation of prompting strategies into a structured taxonomy, cataloguing dozens of techniques (few-shot, chain-of-thought, role framing, meta-prompts, prompt chaining, and modality-specific methods) and mapping them to applications and evaluation metrics. The paper emphasizes reproducibility and recommends best practices clear task framing, exemplar selection, parameter tuning, and automated prompt testing to avoid brittle or dataset-specific prompt designs. Importantly, Schulhoff stresses tooling: prompt libraries, validators, and testing frameworks are necessary to scale prompt engineering beyond ad-hoc experimentation. Technically, the survey clarifies that prompt engineering is not a single skill but a suite of interlocking practices (templating, parameter sweeps, chaining, retrieval integration, and evaluation), each with its own cost/benefit trade-offs. By codifying these methods, the work provides a practical roadmap for practitioners and organizations to professionalize prompt engineering, while also highlighting the knowledge and tooling gaps that may impede adoption across different user groups.

Aldasoro et al. (2024) analyzed demographic controls and finds that the observed gap is not fully explained by age, education, occupation, or privacy preferences, suggesting that other factors such as self-confidence, exposure, or perceived usefulness contribute. For a technical topic like prompt engineering, this work implies that differences in baseline adoption of generative tools will cascade into disparities in acquiring advanced skills: if fewer women engage with GenAI at all, fewer will encounter or learn prompt engineering techniques. The paper's empirical quantification of the gap provides an evidence base for technical interventions (UI design, guided templates, toolkits) aimed at lowering the initial barriers to entry that precede more complex prompt engineering learning. The study thus positions adoption-level differences as a prerequisite issue that shapes who engages with prompt engineering toolchains.

Russo investigated attitudinal dimensions linked to AI adoption and finds that women report higher AI anxiety, lower perceived knowledge, and lower reported usage of AI tools compared to men. The study identifies a negative relationship between AI anxiety and positive attitudes toward AI that partly mediates gender-based differences in adoption. Technically relevant to prompt engineering, the findings suggest that affective and metacognitive factors influence willingness to experiment with model parameters, advanced prompting techniques, and multi-step prompting workflows. From an engineering standpoint, higher anxiety or lower perceived competence may reduce exploratory parameter sweeps, iterative debugging, and engagement with developer-oriented tools activities central to advanced prompt engineering. Russo therefore links psychosocial variables to technical adoption barriers, supporting the argument that inclusive prompt engineering must address both tooling and user confidence dimensions.

Møgelvang (2024) analyzed generative AI chatbot use among higher-education students in Norway and reported that male students engage more frequently across a broader set of applications. Their work highlights usage patterns (editing, idea generation, coding help) and suggests that disciplinary orientation and prior technical exposure mediate differences. Technically, the study implies that differential curricular exposure to coding, API use, and computational thinking contributes to how readily students adopt advanced prompting methods (e.g., chaining prompts into project workflows, integrating retrieval, or fine-tuning). For practitioners creating educational modules or tooling to teach prompt engineering, the paper signals the importance of domain-tailored examples, scaffolded templates, and hands-on labs to equalize opportunity for advanced skill uptake. The MDPI study thus connects observable usage gaps to specific technical affordances and learning environments.

Deloitte (2024) analysis documents trends in generative AI adoption by gender, finding that while adoption among women lagged early on, the gap narrowed quickly through 2024 in certain geographies; however, trust and confidence in using gen-AI for critical tasks remain lower among women. From a technical standpoint, Deloitte recommends product design interventions guided prompts, in-app

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

templates, low-friction parameter controls, and trust signals to accelerate adoption and safe use. The report's industry perspective is useful because it translates technical fixes (simplified UIs, prebuilt prompt templates, adaptive help) into measurable increases in engagement. For prompt engineering, this suggests that democratizing advanced prompting is less about basic model mechanics and more about interface affordances and trust-building: embedded examples, sandboxed experimentation, and transparency about model limitations can lower the barrier to trying advanced prompt engineering techniques.

Santana Jr. et al. provided a rigorous technical comparison of 14 established prompt engineering techniques over 10 software engineering (SE) tasks (such as code generation, bug fixing, question answering) using four different LLMs. They classify techniques along six core dimensions: Zero-Shot, Few-Shot, Thought-Generation, Ensembling, Self-Criticism, and Decomposition. Their experiments show that for tasks requiring complex logic and reasoning (e.g., bug fixing, code synthesis) techniques involving thought generation or decomposition outperform simpler methods, but at the cost of higher token usage and longer runtime. They also examine how linguistic features of prompts (e.g., prompt length, complexity, example selection) correlate with effectiveness. Technically, this work reveals that different prompting methods have distinct trade-offs depending on task type and model architecture, and that efficient engineering of prompts involves managing not just structure but cost, latency, and resource consumption. For users unfamiliar with advanced prompt engineering, this paper provides a guide for selecting methods appropriate for different kinds of software engineering tasks.

Shin, Mohati et al. (2023) compared several prompting strategies basic prompting, in-context learning, and task-specific prompting against 18 fine-tuned smaller/older models on automated software engineering tasks such as code generation, code summarization, and code translation. They find that while task-specific prompting of a large model (GPT-4) can outperform fine-tuned models in some tasks (e.g., comment generation), it fails in others (e.g., certain code generation tasks), often because fine-tuned models are more optimized for specific distributions. They also note that conversational prompting (interactive, user feedback loop) improves performance beyond static prompt templates. The technical insight is that prompt engineering can approximate the performance of fine-tuning in some cases, but not uniformly, and that the user's ability to iterate and give feedback to the model (i.e. human-in-the-loop) is a key technical lever.

Chen et al. investigated how different prompt engineering strategies affect the performance of an LLM on document information extraction tasks. They experiment with variations in prompt templates, inclusion of instruction vs. examples vs. schema, and compare zero-, one-, few-shot prompting modes. Their results show that prompts which combine schema definitions and example annotations significantly improve extraction accuracy, especially in structured documents. They also study how changing the wording, ordering, and clarity of instruction in prompts impacts precision and recall. Technically, this work indicates that prompt design (template structure, example selection, schema incorporation) strongly influences extraction performance, and small changes can yield measurable differences. For those adopting prompt engineering, mastering document-structure aware prompts and example annotation is crucial.

Sant et al. (2024) examined how different prompt structures affect gender bias in machine translation tasks. They assess LLMs vs. NMT models for translations from English to Spanish and Catalan, using standard test sets, and quantify gender bias using WinoMT and other benchmarks. The core technical finding is that certain prompt structures reduce gender bias by up to ~12% compared to more naïve prompting. The work shows that prompt engineering is not only about performance in accuracy and fluency, but also in fairness and bias mitigation. It also underscores that the choice of prompt wording and framing can materially affect biased outputs, meaning that users who understand

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

prompt structure can influence these bias metrics. This has technical importance for adoption, because it suggests users need to understand bias-related prompting practices to get more equitable results.

OBJECTIVES

The rapid advancement of Generative Artificial Intelligence (AI) has transformed digital creation, communication, and problem-solving. Within this evolution, prompt engineering has emerged as a critical skill that determines how effectively users can interact with AI systems to generate accurate, creative, and context-specific outputs. Despite its growing relevance, the adoption and understanding of prompt engineering as a digital competency remain uneven across different user groups, particularly across genders. Existing literature on technology adoption and digital literacy reveals persistent gender-based disparities in exposure, confidence, and participation in emerging technologies. Numerous studies highlight that women, in comparison to men, often face structural, educational, and perceptual barriers that limit their engagement with advanced digital tools. However, while substantial research exists on gender differences in general technology use and AI awareness, there is limited scholarly exploration specifically addressing gender differences in the adoption and practice of prompt engineering within the context of generative AI. Furthermore, theoretical frameworks such as the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) have been extensively applied to explain technology adoption behaviors. Yet, their applicability to the domain of prompt engineering, especially in analyzing gender-based adoption patterns, remains underexplored. Similarly, secondary data sources, industry reports, and academic analyses have not been systematically synthesized to uncover patterns and influencing factors behind gendered adoption trends in this rapidly evolving field. This gap in comprehensive, theory-driven secondary research presents a critical problem. Without a structured understanding of how gender influences awareness, adoption, and skill development in prompt engineering, AI literacy initiatives risk reinforcing existing digital inequalities. Moreover, the lack of consolidated evidence limits policymakers, educators, and organizations in developing gender-inclusive strategies for AI education and workforce readiness. Therefore, this study seeks to synthesize existing research, examine gender-related patterns, and identify barriers and enablers influencing the adoption of prompt engineering. By integrating insights from theoretical models and empirical literature, the study aims to contribute to the development of gender-equitable frameworks for AI literacy, ensuring balanced participation in the emerging ecosystem of generative AI technologies.

The objectives of current research are :

1. To review and synthesize existing literature on the concept, evolution, and importance of prompt engineering within the domain of generative AI.
2. To examine and compare existing studies and reports that discuss gender-based differences in digital technology adoption, AI usage, and related skill acquisition.
3. To analyze secondary data, academic research, and industry surveys to identify observable patterns or trends in the awareness and adoption of prompt engineering across genders.
4. To explore theoretical frameworks (such as the Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology) that explain gender-based variations in technology adoption and their applicability to prompt engineering.
5. To identify critical factors and barriers influencing gender disparities in adopting prompt engineering, as highlighted in previous research and policy documents.
6. To propose strategies and recommendations for enhancing gender-inclusive AI literacy and promoting equitable participation in prompt engineering practices.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

METHODS

The present study adopts a descriptive–analytical research design grounded in the systematic analysis of secondary data. This design has been selected for its suitability in exploring emerging, conceptually rich domains specifically, the evolution and application of prompt engineering within the context of generative artificial intelligence (AI) and the examination of gender-related differences in technology adoption and digital literacy. The descriptive component facilitates the identification and organization of relevant studies, while the analytical aspect enables critical evaluation and synthesis of existing findings and theoretical frameworks. The study does not engage in any form of primary data collection from respondents. Instead, it relies exclusively on a structured literature review, comparative assessment, and conceptual synthesis of published academic, industrial, and policy-based sources. This approach ensures that the investigation remains grounded in credible, peer-reviewed evidence while allowing for a comprehensive understanding of gendered patterns in prompt engineering adoption and practice.

Sources of Data

Given the secondary nature of the present study, all data and information are derived exclusively from previously published and authenticated sources. This reliance on secondary materials ensures that the analysis is grounded in credible, peer-reviewed, and evidence-based literature while maintaining a broad and comprehensive scope. To ensure both depth and academic reliability, the study draws upon five major categories of secondary data sources: academic journals and research papers, books and book chapters, industry reports and white papers, government and policy documents, and online databases and repositories.

Peer-reviewed academic journals and research papers constitute the primary foundation of this study, offering empirically validated insights into the phenomena of interest. Publications from reputed international journals such as *AI & Society*, *Computers in Human Behavior*, *Information Systems Research*, *Gender, Technology and Development*, and *Technology in Society* provide a rigorous understanding of human–AI interaction, gender-based disparities in technology adoption, and the evolving role of prompt engineering as a digital competency. These sources serve as the academic backbone of the analysis, enabling a critical synthesis of existing research findings and theoretical perspectives.

In addition, books and book chapters are utilized to trace the theoretical and conceptual evolution of artificial intelligence, human–computer interaction, and digital literacy. These scholarly texts contribute to building the conceptual framework of the study by offering historical context and comprehensive discussions on gender inclusion in technology. Authoritative academic volumes and edited collections help bridge theoretical discourse with applied understanding, providing the necessary depth to interpret both empirical data and conceptual models related to prompt engineering adoption.

Complementing the academic sources, industry reports and white papers are reviewed to capture current global trends in AI adoption, workforce development, and gender representation in emerging technologies. Reports issued by institutions such as the *McKinsey Global Institute*, *World Economic Forum (WEF)*, *UNESCO*, *OECD*, *PwC*, and *Deloitte Insights* are particularly valuable for understanding practical and organizational perspectives on AI literacy, skills development, and digital inclusion. These documents offer a data-driven view of industry practices and help identify the socio-economic dimensions influencing gendered participation in AI-related fields.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Furthermore, government and policy documents play a crucial role in contextualizing the study within the broader institutional and policy framework. National and international policies issued by bodies such as *NITI Aayog (India)*, the *Ministry of Electronics and Information Technology (MeitY)*, and *UN Women* are examined to evaluate the extent to which public initiatives, digital inclusion programs, and AI literacy missions address gender disparities. These documents also provide insights into how policy-level strategies influence awareness, accessibility, and skill development in the area of generative AI.

Finally, online databases and digital repositories such as *Google Scholar*, *Scopus*, *IEEE Xplore*, *SpringerLink*, and *Statista* are extensively employed to collect recent studies, reports, and statistical data relevant to the topic. The inclusion of these sources ensures that the study incorporates the latest developments, emerging patterns, and evolving academic discourse surrounding gender differences in AI adoption and prompt engineering. Collectively, these diverse yet complementary sources provide a robust foundation for conducting a comprehensive and critical secondary analysis, ensuring both reliability and academic rigor in addressing the research objectives.

Data Collection Procedure

The process of data collection in this study follows a systematic and structured approach appropriate for secondary research. Since the investigation is based exclusively on existing literature, reports, and databases, the primary goal of data collection is to ensure the identification, selection, and extraction of relevant, credible, and high-quality sources that collectively address the research objectives. The procedure is divided into multiple sequential stages beginning with keyword identification and culminating in the extraction and organization of pertinent information for analysis. This multi-stage approach ensures methodological rigor, comprehensiveness, and the inclusion of diverse perspectives across academic, industrial, and policy domains.

The first stage involves keyword identification, which serves as the foundation for targeted literature retrieval. Based on the study's objectives and thematic scope, a set of keywords and key phrases was developed to guide the search process. These included terms such as *"prompt engineering," "generative AI adoption," "gender differences in technology use," "AI literacy," "digital inclusion," "technology acceptance,"* and *"Technology Acceptance Model (TAM)."* These keywords were selected after a preliminary review of the literature to capture the intersection of gender, artificial intelligence, and digital skill acquisition. The use of carefully chosen keywords ensured that the search yielded relevant results while minimizing irrelevant or redundant material. The second stage consisted of the search strategy, where a systematic retrieval of literature was carried out using established academic databases and repositories such as *Google Scholar, Scopus, IEEE Xplore, SpringerLink,* and *ResearchGate*. Boolean search operators were applied to refine and combine the selected keywords for more accurate search results. For instance, expressions like *"gender AND AI adoption"* or *"prompt engineering AND generative AI"* were used to locate literature that directly or indirectly addressed the study's core areas. This strategic search process enabled the inclusion of both scholarly and applied perspectives, thus ensuring a balanced understanding of the phenomenon under investigation. In the third stage, inclusion and exclusion criteria were applied to maintain the quality and relevance of the selected literature. The inclusion criteria comprised studies published between 2015, written in English, and available in peer-reviewed journals or reputable institutional repositories. These studies were required to focus explicitly on artificial intelligence, prompt engineering, or gender-based variations in technology adoption. Conversely, materials such as non-academic blogs, opinion articles, unverified online content, duplicate reports, and literature predating 2015 that did not align with generative AI developments were excluded. The inclusion and exclusion criteria ensured that only the most relevant and contemporary sources were analyzed, aligning the study with current technological and societal contexts.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The fourth stage entailed a screening process, during which all identified studies were meticulously reviewed to determine their relevance and methodological soundness. The screening was conducted in two phases: an initial examination of titles and abstracts followed by a comprehensive full-text review of shortlisted studies. This rigorous screening ensured that each selected publication contributed meaningful theoretical, empirical, or conceptual insights to the study. In addition, priority was given to studies demonstrating robust research designs, transparent methodologies, and clearly articulated findings related to AI adoption or gender disparities. Finally, the fifth stage involved data extraction and organization, a process aimed at systematically documenting relevant information from the selected sources. For each publication, details such as author(s), year of publication, research purpose, methodology, theoretical framework, major findings, and gender-related insights were recorded in a structured coding matrix. This approach facilitated consistency in data management and allowed for effective comparison and thematic synthesis during the analysis phase. The coding matrix also served as a tool for mapping the relationships among concepts such as prompt engineering, technology adoption behavior, and gender-based differences in AI engagement. Throughout the entire data collection procedure, emphasis was placed on maintaining academic transparency and methodological rigor. The systematic process not only ensured the inclusion of comprehensive and credible literature but also laid the groundwork for meaningful thematic and comparative analysis in the subsequent stages of the study. By adhering to these well-defined procedures, the research achieves both reliability and validity in its secondary data foundation, thereby reinforcing the credibility of the study's interpretations and conclusions.

Data Analysis Techniques

The data analysis for this study is entirely based on qualitative interpretation and conceptual synthesis of secondary information derived from academic literature, institutional reports, and policy documents. Given that no primary data were collected, the analysis focuses on identifying, organizing, and interpreting patterns, trends, and theoretical insights within existing research. The analytical process combines qualitative content analysis, thematic synthesis, and comparative analysis, thereby allowing for a nuanced understanding of gender-based variations in the adoption of prompt engineering and related AI competencies. The initial stage of analysis involves a qualitative content analysis, which focuses on systematically reviewing and categorizing information obtained from the selected sources. Each study, report, or document was closely examined to extract relevant data concerning prompt engineering, generative AI adoption, digital literacy, and gender disparities in technology usage. The extracted data were coded into conceptual categories such as *awareness levels*, *adoption behavior*, *skill acquisition*, *perceived usefulness*, and *barriers to participation*. This process facilitated the identification of recurring concepts and relationships among variables, ensuring that the analysis remained both comprehensive and coherent. Following content organization, a thematic analysis was undertaken to derive broader conceptual insights from the literature. This method enabled the grouping of related findings into meaningful themes that directly align with the objectives of the study. The major themes identified include:

1. The evolution and significance of prompt engineering in enhancing human–AI interaction and its growing importance as a digital literacy skill;
2. Gender-based variations in technology and AI adoption behavior, highlighting socio-cultural, educational, and psychological dimensions;
3. The application of theoretical frameworks such as the Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of Technology (UTAUT) in explaining gendered technology adoption patterns;
4. The identification of barriers and enabling factors influencing AI skill acquisition among different genders; and

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

5. The policy and educational implications of fostering inclusive AI literacy to reduce gender disparities in digital engagement.

The thematic synthesis process involved iterative reading, coding, and reorganization of data to ensure that the themes captured the diversity of perspectives presented across the reviewed literature. Each theme was supported by evidence from multiple studies, thereby enhancing the validity and interpretative strength of the findings. Through this process, the study developed an integrated understanding of how gender shapes both the awareness and application of prompt engineering in the broader context of generative AI. To further refine the interpretation, a comparative analysis was conducted to assess convergence and divergence across different studies and data sources. This comparative lens enabled the identification of consistent global patterns as well as contextual differences based on region, profession, or educational background. By juxtaposing academic research with industry and policy reports, the study was able to highlight both theoretical and practical dimensions of gender-based disparities in prompt engineering adoption. This analysis also revealed gaps in existing literature, particularly the limited empirical focus on gendered skill acquisition in prompt design and AI literacy initiatives. Finally, the analysis integrates insights through the application of conceptual frameworks, primarily the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT). These models provide a robust theoretical lens for interpreting the gendered dynamics of technology adoption. Constructs such as *Perceived Usefulness*, *Perceived Ease of Use*, *Social Influence*, and *Facilitating Conditions* were examined to understand how gender may act as a moderating factor influencing the acceptance and practice of prompt engineering. By mapping the reviewed findings onto these frameworks, the study bridges theoretical understanding with observed behavioral patterns, thereby strengthening the explanatory depth of the analysis.

RESULTS

The secondary corpus analysed for this study encompassing peer-reviewed articles, systematic reviews, policy documents, industry reports, platform analytics and practitioner guides provides a rich yet heterogeneous evidence base concerning the nature of prompt engineering and the gendered contours of engagement with generative artificial intelligence (AI). The synthesis below draws together the most salient patterns, maps them to accepted theoretical models of technology acceptance, explicates the principal barriers and enablers, and offers an integrated set of recommendations supported by the literature.

Prompt engineering: a transdisciplinary socio-technical competency

The reviewed literature consistently characterizes prompt engineering as a transdisciplinary socio-technical competency rather than a narrow programming skill (Lee,; McKinsey, 2023). Across pedagogical reviews and practitioner guidance, three interdependent dimensions recur: (a) prompt composition structuring instructions, contextual cues and exemplars to align model outputs with task goals; (b) iterative optimisation employing evaluation, error analysis and refinement cycles to improve prompt efficacy; and (c) contextual framing tailoring prompts to domain conventions, ethical constraints and audience expectations (Lee,; practitioner white papers). This triadic formulation situates prompt engineering at the intersection of communication, domain expertise and an applied understanding of model behaviour.

The practical implication of this hybridity is twofold. First, educational instruments that treat prompt engineering as purely technical (e.g., code-oriented metrics) will systematically under-represent adoption that occurs via communicative or domain-driven pathways. Second, the multiplicity of entry points (no-code interfaces, template-based curricula, applied case studies) expands the set of

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

pedagogical strategies available to foster inclusive uptake (McKinsey, 2023; Coursera, 2024). Accordingly, the literature advocates blended instructional designs that pair conceptual primers on model affordances with scaffolded, task-based labs that permit low-stakes experimentation.

Table 1: Selected literature on prompt engineering and pedagogy

Source (Author / Org, Year)	Type	Principal contribution
Lee	Peer-reviewed review	Conceptualises prompt engineering as socio-technical; recommends blended pedagogy and task labs.
McKinsey (2023)	Industry report	Documents employer demand for generative-AI competencies and growth in micro-credentialing.
Practitioner white papers	Technical guidance	Provide operational templates (few-shot, CoT), evaluation heuristics and no-code tooling recommendations.
Coursera (2024–25)	Platform analytics & playbook	Analyses enrolment patterns and suggests inclusive course designs (mentorship, flexible pacing).

Gendered patterns in engagement, enrolment and practice

A coherent pattern emerges across multiple types of sources: women are under-represented in formal GenAI technical education and AI engineering occupations, enrol at lower rates in many generative-AI technical courses, and report comparatively higher levels of AI-related anxiety and lower self-efficacy (Coursera, 2024–25; UNESCO, 2021; Russo et al.,). Platform analytics commonly report that women constitute approximately 30–33% of enrolments in many GenAI offerings, a figure that mirrors workforce representation estimates from international surveys (Coursera, 2024–25; WEF, 2023). Developer-community metrics (forums, open-source participation) further show male predominance, thereby amplifying informal tacit learning flows that accelerate technical mastery (Stack Overflow summaries; LinkedIn analyses).

Nevertheless, the literature highlights crucial nuances. Where course design includes mentorship, cohort-based learning, contextualised casework and flexible pacing, female enrolment and completion rates improve markedly (Coursera playbook; program evaluations). Moreover, gendered differences are not simply absence vs. presence; they are often differences of orientation and context. Women more frequently engage generative AI in applied, communicative and domain-specific contexts (education, health communication, policy) while men appear more visible in pipeline integration, optimisation and tooling communities. Consequently, measures of adoption that focus exclusively on technical course enrolments or job titles risk under-estimating substantive, practice-based uptake among women.

Table 2: Evidence summary: gendered patterns in GenAI engagement

Evidence source	Type of data	Key gender-relevant finding
Coursera platform analytics (2024–25)	Enrollment/completion data	Women ≈30–33% of GenAI course enrolments; higher completion with cohort/mentorship.
WEF and LinkedIn analyses (2022–23)	Workforce surveys	Women under-represented in AI engineering roles (~25–35% range across reports).
Systematic reviews (e.g., Russo et al.,)	Empirical syntheses	Women report higher AI anxiety and lower self-efficacy in multiple samples.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Developer community metrics	Forum participation data	Male predominance in developer communities; tacit knowledge transmission advantages for men.

Triangulation: awareness, adoption and practice

Direct, gender-disaggregated measurements of prompt-engineering competence such as objective prompt-quality scores or controlled task-performance assessments are conspicuously scarce in the extant literature. Consequently, this study relies on a triangulation of available proxies (platform enrolments, occupational representation, developer-community participation and attitudinal surveys) to infer likely patterns of awareness and adoption. This triangulated approach, while indirect, permits robust, theory-informed inferences by synthesising convergent signals across diverse data streams (WEF, 2023; Coursera, 2024–25; Russo et al.,). First, an exposure differential is evident: men currently experience greater exposure to technical experimentation through participation in developer communities, open-source projects and employer-sponsored learning pathways. Such exposure provides repeated, low-stakes opportunities for iterative trial-and-error and tacit learning, accelerating the acquisition of prompt-engineering heuristics and optimisation techniques (WEF; platform reports). This differential in experiential opportunities helps explain why men appear more represented in technical ranks and in informal knowledge networks that cultivate advanced prompting skills. Second, a motivational divergence emerges across gendered cohorts. The literature suggests that the salience of adoption drivers differs: men's uptake tends to be more strongly correlated with performance expectations and instrumental experimentation that is, learning insofar as it yields demonstrable productivity or career advantages whereas women's uptake is more sensitive to perceptions of ease, contextual relevance and the availability of social supports (TAM/UTAUT extensions; empirical syntheses). In practice, this means that women are likelier to engage when tools and curricula reduce perceived cognitive cost and explicitly demonstrate applied value in domain-relevant settings. Third, the evidence indicates pronounced design sensitivity: pedagogical and product design features that reduce cognitive and administrative friction such as low-barrier no-code interfaces, reusable prompt templates, domain-relevant casework and cohort-based mentorship disproportionately increase female participation and retention (Coursera playbook; programme evaluations). Where these design elements are present, enrolment and completion differentials shrink, indicating that uptake disparities are responsive to instructional architecture rather than fixed dispositions. Taken together, these interlocking inferences underscore that observed enrolment and participation gaps are best understood as the product of structural exposure disparities, affective barriers (e.g., anxiety, self-efficacy) and design choices, rather than immutable preferences. The implication for educators, platform designers and policy makers is clear: interventions that expand exposure, reduce perceived effort, and strengthen social facilitation are likely to produce measurable gains in equitable prompt-engineering adoption.

Theoretical integration: TAM and UTAUT applied to prompt engineering

The Technology Acceptance Model (TAM) (Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003) offer robust and parsimonious heuristics for interpreting how individuals come to adopt prompt engineering practices within generative-AI ecosystems. In this context, Perceived Usefulness / Performance Expectancy captures the degree to which learners, practitioners and employers believe that prompt-engineering skills will generate tangible productivity, creative or economic returns for example, faster content production, higher-quality analytic outputs, or improved decision-support. Where such task-relevant benefits are clearly articulated and demonstrable, adoption is accelerated; conversely, ambiguity about practical gains suppresses motivation to invest time and effort (Davis, 1989; Venkatesh et al., 2003).

Complementing usefulness, Perceived Ease of Use / Effort Expectancy governs initial engagement: prospective learners are more likely to experiment with prompt design when the perceived cognitive and technical costs are low. The secondary evidence indicates that higher levels of AI-related anxiety and lower self-efficacy—patterns more commonly reported among women in multiple studies—reduce exploratory behaviour unless curricula and tools explicitly scaffold early learning (e.g., no-code interfaces, templates, stepwise labs) (Davis, 1989; Russo et al.,). Thus, reducing cognitive friction and presenting incremental learning pathways are critical levers for broadening participation. UTAUT's constructs of Social Influence and Facilitating Conditions further illuminate uptake dynamics in real-world settings. Social influence peer norms, instructor endorsement, and visibility of role models shapes perceived legitimacy and provides motivational impetus; facilitating conditions access to reliable devices, connectivity, time allocation and organisational supports determine whether adoption can be sustained beyond initial exposure. The literature shows that peer communities, mentors and workplace champions materially accelerate practice adoption, while infrastructural and temporal constraints materially limit it (Venkatesh et al., 2003; Coursera, 2024–25). Crucially, empirical extensions of TAM/UTAUT that treat gender as a moderating variable reveal systematic differences in how these constructs influence behavioural intention. Women's intention to adopt novel technologies tends to be more sensitive to effort expectancy and social facilitation that is, perceptions of ease and the presence of supportive social structures whereas men's intentions often respond more strongly to performance expectancy and demonstrable instrumental gains. Applied to prompt engineering, this moderating effect implies that interventions which reduce perceived effort (guided interfaces, scaffolded exercises) and enhance social support (mentorship, cohort learning) will, on average, yield relatively larger increases in uptake among women; at the same time, communicating clear performance benefits and pathways to applied outcomes remains important for motivating all learners. For research and program evaluation, these insights recommend measuring adoption with instruments that capture both perceptual (usefulness, ease) and contextual (social influence, facilitating conditions) variables, and explicitly testing gender-moderation effects to design targeted, evidence-based interventions.

Barriers: structural, pedagogical and perceptual constraints

The literature identifies a set of persistent, interacting barriers structural, pedagogical and perceptual that together constrain equitable participation in prompt engineering. First, digital access and infrastructure gaps (e.g., lower rates of device ownership and reliable internet) materially reduce opportunities for hands-on practice and asynchronous learning, a constraint that disproportionately affects women in low-resource and rural settings (UNESCO, 2021; Gender Digital Divide indices). Second, educational pipeline deficits, notably the under-representation of women in computing, engineering and related disciplines, limit early exposure to model-centric problem framing and reduce familiarity with the conceptual foundations that facilitate advanced prompt optimisation (WEF, 2023; UNESCO, 2021). Third, affective factors including higher reported AI-related anxiety and lower self-efficacy among many women impede exploratory behaviour, lower persistence when confronted with challenging outputs, and reduce the likelihood of engaging in self-directed experimentation with generative models (Russo et al.,). Fourth, time poverty and role constraints (caregiving responsibilities, unpaid labour and competing work commitments) restrict the discretionary hours available for sustained upskilling, making intensive or poorly-timed training programmes inaccessible for many potential learners (platform playbooks; programme evaluations). Fifth, community and mentorship gaps mean that women frequently lack entry to the informal, tacit knowledge networks (developer forums, peer mentorship, hackathons) that accelerate skill acquisition in male-dominated technical spaces, thereby reducing opportunities for knowledge sharing and problem-solving collaboration (developer community analyses). Finally, trust, ethics and safety concerns including anxiety about biased outputs, privacy, and model misuse can discourage adoption when curricula and training do not explicitly address these issues or provide learners with mitigation strategies (policy and ethics reports). Importantly, these barriers do not act in isolation: limited access amplifies time constraints; lack of

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

mentorship exacerbates anxiety; and pedagogical designs that ignore ethics deepen trust deficits. Together, they create a multifactorial impediment to equitable uptake, signalling the need for multilevel interventions that simultaneously address infrastructure, curriculum design, affective support and institutional facilitation.

Table 3: Barriers and representative evidence

Barrier	Description	Representative sources
Digital access	Device/internet disparities reduce practice opportunities	UNESCO (2021); Gender Digital Divide indices
Educational pipeline	Low female enrolment in computing/engineering	UNESCO; WEF reports
AI anxiety	Higher reported anxiety / lower self-efficacy among women	Russo et al., survey syntheses
Time constraints	Care responsibilities limit upskilling time	Coursera playbook; program evaluations
Mentorship gaps	Fewer women in tacit knowledge networks	Developer community analyses
Ethics & trust concerns	Fear of bias and unsafe outputs reduces adoption	Policy briefs (UN Women, World Bank)

Enablers and evidence-based interventions

The literature converges on a coherent set of enablers and evidence-based interventions that address both perceptual (e.g., anxiety, effort expectancy) and structural (e.g., access, time, institutional support) obstacles to equitable participation in prompt engineering. First, scaffolded entry points including no-code interfaces, reusable prompt templates, and guided laboratory exercises have been shown to reduce perceived effort and early task friction, enabling novices to achieve quick, confidence-building successes and thereby lowering the barrier to continued engagement (Coursera, 2024–25; pedagogical guides). Second, contextualised, domain-relevant curricula that foreground concrete use cases (for example, education, public health, or small business workflows) increase perceived usefulness and directly align learning outcomes with learners' occupational incentives, which in turn fosters motivation and uptake among groups who prioritise applied outcomes. Third, cohort-based mentorship programmes replicate the tacit knowledge transmission characteristic of developer communities within intentionally supportive settings; evidence from program evaluations suggests that mentorship and peer cohorts improve persistence, completion rates, and confidence particularly for women and other under-represented learners. Fourth, institutional and policy supports such as subsidised devices and connectivity, time-flexible training allowances, employer partnerships for work-integrated micro-credentials, and routine collection of gender-disaggregated participation metrics create the facilitating conditions necessary for scale and sustainability (WEF, 2023; UNESCO, 2021; Coursera, 2024–25). Finally, embedding ethics and bias literacy into prompt-engineering syllabi addresses trust and safety concerns that commonly deter adoption: teaching learners how to recognise, interrogate, and mitigate biased outputs both increases confidence and promotes responsible, accountable practice (policy and ethics literature). Collectively, these interventions map directly onto key TAM/UTAUT levers reducing effort expectancy, raising perceived usefulness, strengthening social influence, and improving facilitating conditions and therefore offer a theoretically grounded and practically actionable pathway to more gender-inclusive AI literacy.

Table 4: Recommended interventions: rationale and expected effects

Intervention	Rationale	Expected effect

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

No-code guided labs & templates	Lowers technical barrier and effort expectancy	Increased early engagement and confidence
Domain-specific micro-credentials	Raises perceived usefulness for applied roles	Higher uptake in women concentrated in applied sectors
Mentorship & cohort models	Creates social facilitation and tacit learning	Improved retention and completion rates
Subsidised access & time-flexible programs	Addresses infrastructural and time constraints	Broader participation from under-served groups
Ethics/bias modules	Reduces trust barriers; builds critical skills	Greater willingness to adopt and apply prompts responsibly

CONCLUSION

This study has examined the landscape of prompt engineering within generative AI through a focused secondary-research lens, synthesizing academic literature, industry reports, platform analytics and policy documents to explicate how gender shapes awareness, adoption, and practice. The principal contribution is a consolidated, theory-informed account that treats prompt engineering as a transdisciplinary socio-technical competency one that intersects domain expertise, communicative framing and iterative model-oriented problem solving and that highlights how structural conditions, affective factors and pedagogical design jointly determine who gains access to and benefits from this emergent skill. Empirically grounded in triangulated secondary evidence, the study produced three interlocking inferences. First, an exposure differential currently privileges men: greater presence in developer communities, technical roles and employer-sponsored learning pathways increases opportunities for tacit learning and iterative experimentation that accelerate prompt-engineering expertise. Second, there is a motivational divergence in adoption drivers: men's uptake correlates strongly with perceived performance benefits and experimentation, whereas women's uptake is more sensitive to perceived ease of use, contextual relevance and social facilitation (mentorship, cohort support). Third, prompt-engineering adoption is design sensitive instructional and product choices (no-code interfaces, templates, domain-relevant modules, cohort mentorship) disproportionately expand participation among women by lowering cognitive friction and strengthening facilitating conditions. Collectively, these findings indicate that observed gender gaps are not immutable preferences but rather the outcome of differential exposure, affective barriers and design choices that policy and pedagogy can address. Theoretically, the review reinforces and extends classic acceptance frameworks (TAM, UTAUT) by demonstrating their applicability to prompt engineering while underscoring the need to treat gender as an explicit moderating variable. Mapping the secondary evidence onto constructs such as *Perceived Usefulness/Performance Expectancy*, *Perceived Ease of Use/Effort Expectancy*, *Social Influence* and *Facilitating Conditions* clarifies why particular interventions (e.g., scaffolded labs to lower effort; mentorship to strengthen social influence; employer endorsements to augment perceived usefulness) are likely to be effective, and it points to measurement priorities when evaluating interventions across gender groups. From a practical and policy perspective, the synthesis yields a coherent, evidence-aligned strategy for promoting gender-inclusive AI literacy and prompt-engineering participation. At the pedagogical level, curriculum designers should prioritise blended, scaffolded approaches that combine conceptual primers with low-barrier practicums (no-code labs, guided templates), domain-contextualised micro-credentials, and embedded ethics/bias literacy. At the organisational level, employers and training providers should support time-flexible, work-integrated micro-credentials, create mentorship and cohort structures that reproduce tacit learning flows, and codify gender-disaggregated monitoring of enrolment, completion and labour-market outcomes. At the policy level, investments in connectivity and device access, subsidies for targeted training, and incentives for employer partnerships are necessary to create the facilitating conditions that enable equitable scale. These interventions act on TAM/UTAUT levers reducing effort, increasing perceived usefulness, strengthening social influence and improving facilitating conditions and therefore offer a theoretically coherent path to narrowing observed gaps. The study also identified important limitations of the extant evidence and of the present analysis. Most notably, direct, gender-disaggregated measures of prompt-engineering competence (objective prompt quality metrics, controlled task performance, transfer of skill to workplace outcomes) are scarce, which constrains causal claims about skill differentials and limits precise effect-size estimation. Secondary sources are heterogeneous in definitions and reporting practices; platform analytics are often proprietary and geographically skewed toward higher-income contexts; and the inclusion of non-binary and gender-diverse populations in available datasets is minimal. Methodologically, this necessitates caution in generalising quantitative magnitudes and underscores the inference-based nature of the present conclusions. Accordingly, the study proposes an agenda for future research. Key priorities include: (1) the development and validation of objective prompt-quality and task-performance metrics that can be reliably administered across contexts; (2) representative, gender-disaggregated data collection across diverse geographies (including LMICs) and occupational sectors; (3) longitudinal and mixed-methods studies that track learning trajectories, skill persistence and labour-market outcomes following prompt-engineering training; (4) experimental or quasi-experimental evaluations of pedagogical and product interventions (no-code labs, mentorship models, domain micro-credentials); and (5) inclusive data collection that captures non-binary and gender-diverse experiences to ensure interventions do not reproduce exclusion in new forms. Prompt engineering is poised to be a defining competency of the AI era; if left unaddressed, gendered disparities in its adoption risk reproducing and amplifying existing inequalities in digital opportunities and labour market outcomes. However, the evidence assembled indicates clear, actionable levers pedagogical, organisational and policy that, when implemented in combination, can lower barriers, expand access and foster inclusive capability building. Translating these insights into validated programmes and measured outcomes will require coordinated effort across educators, employers, platforms and policymakers, supported by rigorous empirical research that directly measures skill and impact. The task is urgent: to ensure that the benefits of generative AI and the capacities to shape it responsibly are distributed widely and fairly.

REFRENCES

[1] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). *Language models are few-shot learners. Advances in Neural Information Processing Systems, 33*, 1877–1901. https://doi.org/10.48550/arXiv.2005.14165

[2] Wei, J., Wang, X., Schuurmans, D., Le, Q. V., Bosma, M., Xia, F., ... & Zhou, D. (2022). *Chain-of-thought prompting elicits reasoning in large language models. Advances in Neural Information Processing Systems, 35*, 24824–24837. https://doi.org/10.48550/arXiv.2201.11903

[3] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems, 33*, 9459–9474. https://doi.org/10.48550/arXiv.2005.11401

[4] Schulhoff, S. (2024). A systematic survey of prompting techniques. *arXiv*. https://doi.org/10.48550/arXiv.2406.06608

[5] Aldasoro, I. (2024). The gen AI gender gap. *Bank for International Settlements Working Paper No. 1197*. https://doi.org/10.2139/ssrn.3702625

[7] Shin, J., Mohati, S., & others. (2023). Prompt engineering or fine-tuning: An empirical assessment of large language models in automated software engineering tasks. *arXiv*. https://arxiv.org/abs/2310.10508

[9] Sant, A., Escolano, C., Mash, A., De Luca Fornaciari, F., & Melero, M. (2024). Evaluating and mitigating gender bias in machine translation with large language models. *ACL Anthology*. https://aclanthology.org/2024.gebnlp-1.7/

[10] Kepel, D., & Valogianni, K. (2024). Autonomous prompt engineering in large language models. *arXiv*. https://doi.org/10.48550/arXiv.2407.11000

[11] Jin, C., Peng, H., Zhao, S., Wang, Z., Xu, W., Han, L., Zhao, J., Zhong, K., Rajasekaran, S., & Metaxas, D. N. (2024). APEER: Automatic prompt engineering enhances large language model reranking. *arXiv*. https://doi.org/10.48550/arXiv.2406.14449

[12] Shah, C. (2024). From prompt engineering to prompt science with human in the loop. *arXiv*. https://doi.org/10.48550/arXiv.2401.04122

[13] Vatsal, S., & Dubey, H. (2024). A survey of prompt engineering methods in large language models for different NLP tasks. *arXiv*. https://doi.org/10.48550/arXiv.2407.12994

[15] Jin, F., Liu, Y., & Tan, Y. (2024). A survey of automatic prompt optimization with instruction tuning. *arXiv*. https://doi.org/10.48550/arXiv.2502.18746

[16] Luccioni, S., Jernite, Y., & Strubell, E. (2024). Green prompting: Mitigating environmental impact of large language models through efficient prompt design. *arXiv*. https://doi.org/10.48550/arXiv.2503.10666

[17] Sherson, J., & Vinchon, F. (2024). Facilitating human feedback for GenAI prompt optimization. *arXiv*. https://doi.org/10.48550/arXiv.2404.15304